ASSESSMENT OF REPORTED DIFFERENCES BETWEEN EXPENDITURES AND LOW INCOMES IN THE U.S. CONSUMER EXPENDITURE SURVEY

J. L. Eltinge, Texas A&M University, I. S. Yansaneh, Westat, and G. D. Paulin, BLS

J. L. Eltinge, Department of Statistics, Texas A&M University, College Station, TX 77843

Abstract. This paper presents some exploratory analyses of differences between income and expenditure values reported by single-person consumer units in the U.S. Consumer Expenditure Interview Survey. Simple descriptive statistics suggest the potential for serious concern regarding these differences. For example, for units in low reported-income groups, mean reported income is substantially lower than mean reported expenditures. Four complementary analyses are used to assess such differences. (1) Comparison of the distributions of reported incomes and expenditures. (2) Similar comparisons for specialized demographic groups (e.g., retired, student, or self-employed), and groups of units with special reporting patterns (e.g., high or low first-interview income reports). (3) Estimation of a model for the propensity to report income that is substantially lower than reported expenditures. (4) Use of measurement error models and related sensitivity analyses.

Key words. Errors-in-variables, Exploratory analysis, Measurement error, Reliability data, Reported dissaving, Underreporting.

1. Introduction

The U.S. Consumer Expenditure Interview Survey is a major federal household survey sponsored by the Bureau of Labor Statistics. It is one of the few large-scale surveys that collect detailed data on both income and expenditures at a household, or "consumer unit" level. Consequently, this survey is a potentially rich source of information on certain relationships among expenditures, income, and related social and demographic variables. However, a major potential problem in using these data is that many interviewed consumer units report expenditures that are substantially larger than reported income values. This phenomenon is especially noticeable for consumer units with relatively low reported incomes. For example, Bureau of Labor Statistics (1991) reported that in 1990, complete income reporters classified in the lowest estimated quintile of income had an estimated mean reported *before* tax income of $5637 and an estimated mean reported expenditure of $12,908. This phenomenon is sometimes described informally as "reported dissaving."

This reported dissaving appears to involve some consumer units that have true "dissaving" and some consumer units that underreport income or incur other forms of measurement error. For some subpopulations, one may expect true dissaving to occur. For example, college students may use money from student loans or savings to pay for some expenditures. Similarly, one may expect some retired workers to draw down savings. In addition, some self-employed workers and other persons with highly variable incomes may have spending patterns that reflect some "average" of previous, current, and anticipated future income, so that in low-income periods, such persons would also report expenditures that are higher than their current incomes. Finally, in a given quarter, a consumer unit may report "true dissaving" due to an unusually large expenditure, e.g., a down payment for an automobile.

For other groups, this "reported dissaving" may be less easily explained on substantive grounds, and may suggest the presence of substantial measurement errors in reported incomes, reported expenditures, or both. As with many other surveys, measurement error problems in the present case may arise from several sources, e.g., simple recall error, questionnaire design, and the characteristics and perceptions of the respondent and interviewer. However, for the Consumer Expenditure Interview Survey, there are several reasons to expect that substantial measurement errors are generally more common in the reported income data than in the reported expenditure data. For example, work with questionnaire design has focused primarily on obtaining high-quality expenditure data. In addition, respondents' concerns about confidentiality tend to be stronger for income data than for expenditure data, so intentional underreporting of income is expected to be a more severe problem than either intentional underreporting or overreporting of expenditures.

The remainder of this paper examines the extent to which data internal to the Consumer Expenditure Survey permit assessment of the related issues of "dissaving" subpopulations and measurement error. Section 2 gives some background on the Consumer Expenditure Interview Survey, describes the data under consideration and reviews the parameter estimation and variance estimation methodology used in this paper. Section 3 explores the extent to which "reported dissaving" may be associated with specified groups or with poor data quality. The results of Section 3 suggest that for one baseline subpopulation, a substantial part of reported dissaving may be attributable to measurement error. Section 4 considers some implications of this result for errors-in-variables regression of consumer-unit expenditures on reported incomes and on other variables. Departures from standard errors-in-variables methodology follow from use of a data-quality indicator. In addition, limitations on replicate-observation-based identifying information lead to a special errors-in-variables sensitivity analysis. Section 5 summarizes the results of this paper and discusses some methodological alternatives.

2. Data and Estimation Methodology

2.1 Variables of Principal Interest

Table 1 lists the variables of principal interest in this study. Some background is as follows. In the Consumer Expenditure Interview Survey, selected sample "consumer units," roughly equivalent to households, are asked to participate in a total of five interviews. First-interview items include demographic and bounding variables, and, in grouped form, total income for the consumer unit in the preceding twelve months. A resulting classification variable CUINCQ2 will be used in Section 3. The second through fifth interviews are carried out at three month intervals; in each, the consumer unit is asked to report expenditures in the preceding three months. The resulting total expenditure reports for consumer unit i are denoted ERANKMTH_{2i} through ERANKMTH_{5i}, respectively, and the sum of these four reported expenditure amounts is denoted ERANKSUM_i.

Also, in the second and fifth interviews, each consumer unit is asked to report its total before-tax income for the preceding twelve months; the resulting income reports will be denoted FINCBTAX_{2i} and FINCBTAX_{5i}, respectively. Thus, for consumer unit i, FINCBTAX_{5i} is a report of income for the same twelve months covered by ERANKSUM_i. The work in Section 3 will use the differences of logarithms $L_i = \ln(\text{FINCBTAX}_{5i})$ - $\ln(\text{ERANKSUM}_i)$ and the relative differences $R_i = (\text{FINCBTAX}_{5i} - \text{ERANKSUM}_i)/\text{ERANKSUM}_i$.

It should be emphasized that the FINCBTAX variables are defined to represent income *before* taxes. Consequently, one would in general expect FINCBTAX_{5i} to be somewhat larger than ERANKSUM_i, excluding the effects of debt or savings. Some income-tax questions are included in the Consumer Expenditure Interview Survey. However, there is a fairly strong basis for concern regarding apparent underreporting or nonreporting of the relevant tax figures. For example, in 1992 only 56 percent of complete income reporters reported any federal tax payments. In addition, for those units that did report federal tax payments, the mean reported tax amount was less than seven percent of the mean reported income. Consequently, after-tax income variables will not be considered further here. Similarly, it was decided not to use asset and liability data, due to problems with apparent underreporting or nonreporting.

The interviewer records provided some information relevant to assessment of data quality. For the present study, the data-quality variable of principal interest is RECORD_{5i}, which indicates the extent to which consumer unit i used records in answering questions in the fifth interview; RECORD_{5i} takes integer values from 1 (records always used) to 6 (records never used). A reasonable initial conjecture is that units with lower values of RECORD_{5i} may have fewer problems with measurement error in income and other variables reported in the fifth interview. Consequently, Section 3 below will consider the extent to which units with different RECORD_{5i} values show different relationships between FINCBTAX_{5i} and ERANKSUM_i.

2.2 Data Used

The present study is restricted to single-person consumer units interviewed in 1988 through 1992. The restriction to single-person units was chosen to avoid some specific potential measurement error problems associated with proxy reporting, to avoid some definitional issues associated with changes in consumer unit membership across interviews, and to provide some simplification in data management.

All analyses presented here used only data from the 2329 units that were "stringent income reporters." A stringent income reporter is a consumer unit that did not give a refusal or "don't know" answer to any income questions in either the second or fifth interviews. Note that this definition excludes any consumer unit that did not participate in one or both of the second or fifth interviews. For a similar restriction to a set of "stringent" income reporters in a single interview, see Garner and Blanciforti (1994). In this work, units which are not stringent income reporters are called "incomplete income reporters."

2.3 Estimation Methodology

All point estimates reported here were computed using weights that were adjusted for incomplete income reporting. The main steps in this adjustment were as follows. First, logistic regression models were fit for the "stringent income reporter" indicators, using the fifth-interview FINLWT21 weights as the initial survey weights. Second, following the general strategy of Little (1986) (see also Rosenbaum and Rubin (1983) and Czajka et al. (1992)), as implemented in Eltinge and Yansaneh (1993), adjustment cells based on the estimated stringent-reporter probabilities were constructed, and weight-adjustment factors were computed accordingly. The resulting survey weights were equal to the product of the weight-adjustment factors multiplied by the original FINLWT21 weights. See Eltinge and Yansaneh (1993) for a detailed discussion of the cell-construction and adjustment-factor methodology.

All standard errors reported in this paper are based on variance estimates computed through the balanced repeated replication method as implemented for the Consumer Expenditure Interview Survey. This balanced repeated replication method is intended to account both for variability due to the original sampling design and for variability due to nonresponse and subsequent weighting adjustments. The standard errors in Section 4 also use an additional approximation. For some general background on balanced repeated replication, see, e.g., Wolter (1985, Chapter 3). In the present work, the balanced repeated replication computations were based on a set of 44 replicate weights computed in a manner parallel to the computation of the standard FINLWT21 weight. For the present weighting adjustment work with the stringent income reporters, each of the 44 replicate weights was adjusted by: (1) fitting a weighted logistic regression model for "stringent income reporter" indicators using one

Table 1: Variables and Classifications of Principal Interest

Variable	Description
CUINCQ2	Self-reported classification of units according to income in the twelve months prededing Interview 1 (1: loss; 2: [0, 6000); 3: [6000, 10,000); 4: [10,000, 20,000); 5: [20,000, 35,000) 6: > 35,000)
$FINCBTAX_2$	Reported income before taxes for the twelve months preceding Interview 2
$FINCBTAX_5$	Reported income before taxes for the twelve months preceding Interview 5
$ERANKMTH_2$	Total reported expenditures for the three months preceding Interview 2
$ERANKMTH_3$	Total reported expenditures for the three months preceding Interview 3
$ERANKMTH_4$	Total reported expenditures for the three months preceding Interview 4
$ERANKMTH_5$	Total reported expenditures for the three months preceding Interview 5
ERANKSUM	Sum of $ERANKMTH_2$, $ERANKMTH_3$, $ERANKMTH_4$, $ERANKMTH_5$
Age	Age of person in consumer unit
Any College	Person reported some college, college degree, or postgraduate education
Black	Person is black
Female	Person is female
Urban	Person is in an urban area
Urban, Midwest	Person is in an urban area of the midwest
Urban, South	Person is in an urban area of the south
Urban, West	Person is in an urban area of the west
Renter	Person rents residence
Owner, No Mortgage	Person owns residence, with no mortgage
Housing Assistance	Person reports housing assistance
Income Assistance, Int. 5	Person reports income assistance in interview 5
$RECORD_5$	Classification of units according to record use in Interview 5 ($RECORD_5 = 1$ "Always uses records" to $RECORD_5 = 6$ "Never uses records")
High Record Use, Int. 5	Had $RECORD_5 = 1$ or 2 or 3

Table 2: Partition of the Singles Population into Subpopulations

Subpopulation	Description
1	Units that are retired
2	Units (not in subpopulation 1) that are students
3	Units (not in subpopulations 1 or 2) that are self-employed
4	Units that are not in subpopulations 1 through 3 ("baseline" subpopulation)

Table 3: Descriptive Statistics for ln(FINCBTAX$_5$) − ln(ERANKSUM), by Subpopulation Membership

Subpopulation	$\bar{x}$	$s(\bar{x})$	$q_{.25}$	$q_{.50}$	$q_{.75}$	Proportion
All (1 - 4)	0.0586	0.0161	-0.1418	0.1535	0.3951	
Retired (1)	0.0331	0.0252	-0.2194	0.0493	0.2910	0.300
Student (2)	-0.2277	0.0724	-0.5478	-0.0458	0.2855	0.110
Self-Employed (3)	-0.0275	0.1006	-0.3187	0.0204	0.4504	0.049
Baseline (4)	0.1389		-0.0267	0.2444	0.4419	0.541

Table 4: Descriptive Statistics for ln(FINCBTAX$_5$) − ln(ERANKSUM), by Previous-Income Classification CUINCQ2, Baseline Subpopulation

Income Group	$\bar{x}$	$q_{.25}$	$q_{.50}$	$q_{.75}$	Proportion
[0, 6000)	-0.3002	-0.5398	-0.1134	0.1448	0.104
[6000, 10,000)	0.0233	-0.1613	0.0574	0.3055	0.087
[10,000, 20,000)	0.1327	-0.0385	0.1993	0.3981	0.296
[20,000, 35,000)	0.2290	0.0709	0.3049	0.4741	0.337
> 35,000	0.3063	0.2066	0.3861	0.5640	0.176

of the 44 original sets of replicate weights; (2) construction of adjustment cells and weighting adjustment factors based only on the information from (1); and (3) modification of the relevant set of replicate weights, using the adjustment factors constructed in (2). Steps (1) through (3) were repeated for each of the 44 original sets of replicate weights; the result was a new set of 44 replicate weights adjusted for the "stringent income reporting" restriction.

It should be emphasized that the present work uses one of several possible approaches to point estimation and variance estimation for the parameters of interest. Section 5 discusses briefly some alternative estimation methods that could also be used for the issues considered here. Analyses using these alternative methods are also of interest, but are beyond the scope of the present work.

3. Exploratory Analysis of Reported Income and Expenditures

3.1 Descriptive Statistics for Subpopulations

As noted in Section 1, it was expected that certain subpopulations were more likely to report "true dissaving;" and that some other reports of dissaving might be associated with underreporting of income or overreporting of expenditures. To study these conjectures further, we partitioned the singles population into the four subpopulations indicated in Table 2. Note that the first three subpopulations contain units for which reported dissaving may not be surprising, due to employment status. Subpopulation 4, henceforth called the "baseline subpopulation," consists of the remaining units in the singles population. Within the baseline group, reported dissaving may more readily be attributed to measurement error associated with underreporting of income, overreporting of expenditures, or both.

Table 3 presents descriptive statistics for the variable L_i in subpopulations 1 through 4, including the estimated subpopulation mean, the standard error of the mean, the estimated first through third quartiles, and the estimated proportion of the singles population falling into the indicated subpopulation. Several points are worth noting. First, all four of the subpopulations had estimated first quartiles less than zero. Thus, "reported dissaving" is a fairly common phenomenon within each of the subpopulations under consideration. Second, the student subpopulation had a mean estimate substantially below zero, and also had a negative median estimate. The retired subpopulation had mean and median estimates that were positive, but less than the corresponding estimates for the full population or the baseline subpopulation. The self-employed subpopulation had a small negative mean estimate, but also had a relatively high estimated third quartile. Similar results, not detailed here, were obtained for R_i and other related variables.

Table 5: Logistic Regression Coefficient Estimates and Standard Errors for δ_γ "Dissaving" Indicators

γ	0.00	0.00	0.20	0.20	0.50	0.50
Proportion	0.269		0.139		0.064	
Regressor	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.
Intercept	-2.197	(0.855)	-3.27	(1.20)	-4.64	(1.74)
Age	0.0580	(0.0381)	0.0629	(0.0542)	0.1125	(0.0788)
Age^2	-0.000459	(0.000389)	-0.000484	(0.000547)	-0.001088	(0.000798)
Any College	-0.315	(0.228)	-0.392	(.326)	-0.037	(0.336)
Black	0.228	(0.221)	0.572	(0.298)	1.015	(0.470)
Female	0.310	(0.144)	0.042	(0.229)	-0.267	(0.332)
Urban	-0.516	(0.327)	-0.506	(0.281)	-0.886	(0.500)
Urban, Midwest	0.018	(0.313)	0.023	(0.311)	0.044	(0.515)
Urban, South	0.457	(0.249)	0.716	(0.253)	0.388	(0.484)
Urban, West	0.128	(0.304)	0.295	(0.205)	0.090	(0.483)
High Record Use, Int. 5	-0.483	(0.171)	-0.459	(0.258)	-0.482	(0.441)
Renter	0.0151	(0.201)	-0.104	(0.330)	-0.121	(0.360)
Owner, No Mortgage	-0.283	(0.259)	-0.301	(0.342)	-0.106	(0.476)
Housing Assistance	0.074	(0.475)	0.182	(0.566)	1.074	(0.682)
Income Assistance, Int. 5	0.889	(0.277)	1.146	(0.294)	0.464	(0.466)

Table 6: Descriptive Statistics for $\ln(\text{FINCBTAX}_5) - \ln(\text{ERANKSUM})$, by Record Use Categories within the Baseline Subpopulation

Subpopulation	$\bar{x}$	$s(\bar{x})$	$q_{.25}$	$q_{.50}$	$q_{.75}$	Proportion
High Record Use (1 - 3)	0.2244	0.0259	0.0463	0.2825	0.4821	0.368
Low Record Use (4 - 9)	0.0890	0.0307	-0.0522	0.2220	0.4119	0.632

The results in Table 3 are generally consistent with the idea that some of the "reported dissaving" phenomenon is associated with the retired and student groups, while the role of self-employed persons appears to be more complex. Consequently, subsequent analyses in Sections 3.2, 3.3 and 4 will focus on the expanded baseline subpopulation 4. Additional work, which is not detailed here, was carried out with units in subpopulation 4 that had especially volatile income or expenditure reports. However, these additional comparisons offered somewhat less insight than the results reported here, and will not be considered further.

In addition, recall from Section 1 that there was some concern that "reported dissaving" appeared to be especially marked in low-reported-income groups. Routine calculations suggest that this phenomenon could in principle be attributable to simple volatility of income and expenditures, especially if the same income variable is used for income-group classification and for the calculation of L_i. To study this further, Table 4 presents summary statistics for units in subpopulation 4 classified by the first-interview variable CUINCQ2. Note the pronounced pattern of progressively higher mean and quantile estimates in the higher first-interview income groups.

3.2 Logistic Regression Models for Varying Levels of Reported Dissaving

3.2.1 Varying Levels of Reported Dissaving

The preceding subsection examined the distributions of the differences between the logarithms of income and expenditure values in subpopulations 1 through 4. For the baseline subpopulation 4, a complementary approach is to construct models for the probability that a unit exhibits a certain substantial degree of reported dissaving.

Specifically, for $\gamma \in [0, 1)$ and for each consumer unit i, define the indicator variables

$$\delta_{\gamma,i} = \{1 \quad \text{if CU } i \text{ has } R_i < -\gamma, \text{ and } 0 \text{ otherwise}\}$$

Thus, $\delta_{\gamma,i}$ identifies consumer units that have a "reported dissaving" greater than $\gamma \times 100\%$ of reported expenditures. An associated logistic regression model for $\delta_{\gamma,i}$ is then

$$\ln\{P(\delta_{\gamma,i} = 1)/P(\delta_{\gamma,i} = 0)\} = X_i\theta_\gamma$$

where X_i is a vector of demographic and related variables observed for consumer unit i, and θ_γ is the associated vector of logistic regression coefficients.

Parameter estimates and associated standard errors for θ_γ were computed as follows. First, the dataset was restricted to stringent income reporters in subpopulation 4. Second, for this dataset, point estimates $\hat{\theta}_\gamma$ were computed using standard logistic regression with weights equal to the stringent-reporter adjusted weights defined in Section 2. Third, the estimated covariance matrices for $\hat{\theta}_\gamma$ were computed using the general balanced repeated replication method described in Section 2.

3.2.2 Estimation Results

Logistic regression coefficient estimates and standard errors for the δ_γ "reported dissaving" indicators were computed for $\gamma = 0.0(0.05)0.5$. Due to limited space, Table 5 reports results only for $\gamma = 0.0, 0.2$ and 0.5. The first row of Table 5 lists the value of γ under consideration, and the second row reports the estimated proportion of units with R_i less than $-\gamma$. In keeping with standard cautionary remarks about interpretation of multiple hypothesis tests, comparison of point estimates and standard errors across the values of γ should be viewed as an exploratory device, rather than as a set of formal hypothesis tests. Conditional on this cautionary remark, one may note especially the negative coefficients for the variable, "High Record Use" for low or moderate levels of γ; this is consistent with the suggestion that low record use, and possible resulting measurement errors, may be associated with some degree of reported dissaving. Other notable results include the positive coefficients for the variable "Female" at low levels of γ, "Income Assistance" at low or moderate levels of γ, and "Black" at moderate or high levels of γ. In addition, note the negative estimated coefficients for "Urban" and the positive estimated coefficients for "Urban, South." By contrast, other demographic variables, e.g., age, education or housing tenure, were found to be of less interest in this logistic regression setting.

3.3 Comparison of High- and Low-Record-Use Groups

To study further the relationship between record use and reported dissaving, Table 6 compares the distributions of L_i in the high- and low-record-use groups in subpopulation 4. Note especially that the means and quantiles for L_i are substantially higher for the high-record-use group. Thus, Table 6 again is consistent with the suggestion that some of the "reported dissaving" phenomenon may be associated with measurement error. In work not detailed here, more refined comparisons across individual RECORD_{5i} categories were not markedly more informative than that provided by the high- and low-record-use grouping.

4. Errors-in-Variables Sensitivity Analysis

Section 3 suggested that some of the reported "dissaving" observed in the Consumer Expenditure Survey may be attributable to measurement error, especially underreporting of income. Such measurement errors would have important practical implications in several areas, including the estimation of regression models for the relationship between expenditures and income. In general, the presence of measurement error in reported expenditure and income values can produce serious biases in standard regression coefficient estimators. Consequently, it is important to consider alternative estimation methods that account for this "errors-in-variables" bias. However, some standard errors-in-variables estimation methods rely on restrictive assumptions (e.g., the assumption of measurement error means equal to zero) that are not appropriate for the Consumer Expenditure Survey dataset. In addition, standard errors-in-variables methods generally use model identification information (e.g., independent replicated observations of the same unit) not available from the CE dataset. Subsection 4.1 presents some modified errors-in-variables methods that account for the nonzero-error-mean problem and the limited-replication problem. Subsection 4.2 applies the resulting sensitivity analysis to estimation of coefficients for the regression of ln(ERANKSUM$_i$) on ln(FINCBTAX$_{5i}$) and other explanatory variables.

4.1 Modified Errors-in-Variables Methods

4.1.1 Standard Models and Methods

Adapting notation from Fuller (1987, 1991), a standard approach to errors-in-variable estimation is as follows. Suppose for the moment that one observes the vector $Z_i = (Y_i, X_i)$ for each unit i in a population of size N. (Section 4.2 below will consider an extension using weighted estimation based on sample units.) Assume that

$$X_i = x_i + u_i$$

where x_i is a k-dimensional vector of regressor "true values" for unit i and u_i is the corresponding vector of measurement errors. Assume also that the x_i values have common mean μ_x and covariance matrix Σ_{xx}, and that

$$Y_i = x_i\beta + e_i$$

where β is the regression coefficient vector of principal interest. The error term e_i reflects the combined effects of measurement error and equation error in the Y_i variable, and is assumed to have mean zero and variance σ_{qq}. Standard errors-in-variables methodology generally uses the assumption that the measurement error vectors u_i have mean zero and covariance matrix Σ_{uu}. In addition, model identification information is generally provided by an estimator $\hat{\Sigma}_{uu}$ computed from auxiliary data, e.g., independent replicated observations of the same "true value" z_i. See, e.g., Fuller (1987) or Carroll and Stefanski (1990) for detailed discussions of the use of replicates in errors-in-variables estimation. To simplify the present discussion, we also will assume that the errors e_i and u_i are uncorrelated.

Under these assumptions, the matrix of uncorrected second moments, $M_{ZZ} = N^{-1}\sum_{i=1}^{N} Z_i'Z_i$ has expectation equal to

$$(\beta, I_k)'(\mu_x'\mu_x + \Sigma_{xx})(\beta, I_k) + \Sigma_{ee} + \Sigma_{aa}$$

where $\Sigma_{ee} = \mathrm{diag}(\sigma_{ee}, 0)$ and $\Sigma_{aa} = \mathrm{diag}(0, \Sigma_{uu})$. A simple method-of-moments estimator of β is then

$$\hat{\beta} = \hat{M}_{xx}^{-1}\hat{M}_{xy} \qquad (4.1)$$

where $\hat{M}_{xx} = M_{XX} - \hat{\Sigma}_{uu}$ and $\hat{M}_{xy} = M_{XY}$, and M_{XX} and M_{XY} are the indicated submatrices of M_{ZZ}.

4.1.2 Modified Methods

Two notable problems arise in considering the use of estimator (4.1) to perform errors-in-variables regression of ln(ERANKSUM$_i$) on ln(FINCBTAX$_{5i}$) and other explanatory variables. First, in light of the results of Section 3, it is not appropriate to assume that all units have measurement error means equal to zero. Second, the data available do not include independent replicated observations for the same true values x_i, so a standard replicate-based estimator $\hat{\Sigma}_{uu}$ is not available.

In considering both of these problems, we will restrict attention to the problem of measurement errors in the regressor ln(FINCBTAX$_{5i}$). In addition, we will assume that the other regressors considered have zero measurement error, and that the ln(ERANKSUM$_i$) variables have mean-zero measurement errors which are not correlated with the regressors or with the measurement errors in the variables ln(FINCBTAX$_{5i}$). These assumptions are not expected to be perfectly true, but do allow us to focus on the dominant problem of measurement errors in the regressor ln(FINCBTAX$_{5i}$).

To address the first problem of nonzero measurement error means, assume that the measurement errors in ln(FINCBTAX$_{5i}$) have a mean of zero for units in the high-record-use group and a possibly nonzero mean for units in the low-record-use group. Define μ_u to be the resulting weighted average of these two means, and note that M_{ZZ} then has an expectation equal to

$$(\beta, I_k)(\mu_x'\mu_x + \Sigma_{xx})(\beta, I_k) + \Sigma_{ee} + (0, \mu_u)'(0, \mu_u)$$
$$+(\beta, I_k)\mu_x'(0, \mu_u) + (0, \mu_u)'\mu_x(\beta, I_k)' + \Sigma_{aa}$$

and the corresponding method-of-moments estimator of β is again given by (4.1), with $\hat{M}_{xx}$ and $\hat{M}_{xy}$ adjusted accordingly. In that adjustment, the principal problem is the estimation of μ_u. Although the high-record-use group is assumed to have a measurement error mean equal to zero, direct use of the difference between the means of the high-use and low-use groups would not be an appropriate estimator of the measurement error mean for the low-use group, due to possible differences in the means of the true x values for the two groups. Instead, it was assumed that the two groups had the same coefficient vector in the regression of the true ln(FINCBTAX$_{5i}$) on the other regressors under consideration, $X_{(2)}$, say. This coefficient vector was estimated from units in the high-record-use group. The resulting residual terms ln(FINCBTAX$_{5i}$) $- \hat{\beta}_{(2)}X_{(2)i}$ were computed for each unit in the low-record-use group. Under the assumptions stated above, the weighted mean of these residuals is an approximately unbiased estimator of the measurement error mean for the low-record-use group, and the corresponding weighted average of this estimated mean with zero gives an approximately unbiased estimator of μ_u.

To address the second problem, involving lack of pure replicate observations, consider the model,

$$g_i = \ln(\mathrm{FINCBTAX}_{5i}) - \ln(\mathrm{FINCBTAX}_{2i})$$
$$= u_i - f_i + d_i$$

where u_i and f_i are the measurement errors in ln(FINCBTAX$_{5i}$) and ln(FINCBTAX$_{2i}$), respectively, and d_i is the difference between the true fifth- and second-interview FINCBTAX values. In addition, let $\hat{s}_{gg}$ be the weighted sample variance of the g_i terms. Under the excessively simplistic assumptions that d_i is constant for all units and that u_i and f_i are independent and have the same variance, $\hat{s}_{gg}$ would give an approximately unbiased estimator of $2\sigma_{uu}$. Since the latter set of assumptions is not realistic for the present problem, we instead used the following sensitivity analysis approach. Assume for the moment that $\lambda\hat{s}_{gg}$ has an expectation approximately equal to σ_{uu}, where λ is some nonnegative number. Note that $\lambda = 0.5$ corresponds to the simplified model described above, and that $\lambda = 0$ corresponds to the case $\sigma_{uu} = 0$. If λ were known, then $\lambda\hat{s}_{gg}$ would be an approximately unbiased estimator of σ_{uu}. Consequently, substitution into expression (4.1) and use of the above-described modifications for nonzero measurement error means produces an appropriate method-of-moments estimator for β. Since λ is unknown, computation of $\hat{\beta}$ across a range of plausible values for λ gives instead an indication of the sensitivity of the point estimates $\hat{\beta}$ to the magnitude of the measurement error variance σ_{uu}. In general, sensitivity analyses are relatively common in the errors-in-variables literature, and reflect the fact that many datasets subject to substantial measurement error problems do not contain enough model identification information to produce a consistent estimator of β under a realistic measurement error model. For some examples of errors-in-variables sensitivity analyses developed in somewhat different settings, see, e.g., Kalman (1982), Klepper and Leamer (1984) and references cited therein.

4.2 Application of the Modified Methods

The preceding modified errors-in-variables methods were applied to the income and expenditure data described in Sections 2 and 3. The response variable was ln(ERANKSUM$_i$), and the predictors were ln(FINCBTAX$_{5i}$) and the other variables listed in Table 7. The predictors were selected after preliminary screening of other regressors not discussed in detail here. The results in Table 7 are presented to illustrate the proposed errors-in-variables methodology; due to limitations on available data, this analysis is not intended to address fully all relevant economic questions, e.g., the relationship between expenditures and income over a large number of years. In addition, it should be noted that use of before-tax income as a predictor is somewhat nonstandard, but was necessary due to the problems with tax data described in Section 2; see, e.g., Sawtelle (1993) and Gillingham and Hagemann (1983) for other examples of the use of before-tax income as a predictor variable.

The data used in the estimation work were from baseline subpopulation defined in Table 2, with weights as described in Section 2. Table 7 reports coefficient vector estimates estimates and standard errors separately for the values $\lambda = 0, 0.1, 0.2, 0.3, 0.4$ and 0.5. As one would expect, the coefficient for ln(FINCBTAX$_{5i}$) showed a considerable amount of sensitivity to the choice of λ, but other coefficients, e.g., for the "renter" and "housing assistance" regressors also showed considerable sensitivity. Also, as expected, the standard errors are generally increasing in λ, and the combined changes in coefficient estimates and standard errors have a substantial effect on the identification of significant regressors.

5. Discussion

5.1 Review of Results

Exploration of the "reported dissaving" problem for single-person consumer units in the Consumer Expenditure Interview Survey has led to two general sets of results. First, the data analyses presented in Section 3 support the conjecture that some of the observed "reported dissaving" can be attributed to subpopulations for which this pattern is not surprising, e.g., students or retired persons. Second, differences in reporting patterns for the high- and low-record use groups indicates that serious consideration of measurement error issues is also warranted. The specific mean-adjustment and sensitivity-analysis results presented in Section 4 suggest that these measurement error issues have a substantial effect on coefficient estimates and error variance estimates for the present problem, and also illustrate some potential limitations in the availability of model identification information in the present problem.

5.2 Alternative Approaches

The present paper has used one of several possible approaches to point estimation and variance estimation with the data available. Alternative analysis methods are also worth considering. For example, the errors-in-variables sensitivity analysis described in Section 4 can be adapted to other estimation methods, e.g., the general estimating-equation methods of Carroll and Stefanski (1990) or the nonlinear least squares fitting methods of Browne (1984). In addition, one could consider other forms of model identification information, e.g., instrumental variables.

More generally, the estimation methods used with the logistic regression work in Section 3.2 and the errors-in-variables work in Section 4 may be viewed as variants on pseudo-maximum likelihood approaches developed previously. For some general background on pseudo-maximum likelihood methods, see, e.g., Skinner (1989) and references cited therein. Alternatives to the analysis methods used here would involve different approaches to use of survey weights, to variance estimation, and to incomplete-data adjustments. For example, there is a broad range of views regarding the appropriate use of survey weights in regression and other analyses of survey data; Little (1991) discusses several viewpoints. Many of these approaches extend naturally from regression to errors-in-variables estimation, with some additional complexities associated with model identification information.

In addition, the present paper has addressed

Table 7: Errors in Variables Regression Coefficient Estimates: Sensitivity Analysis for the Variance Ratios
$\lambda = 0, 0.1, 0.2, 0.3, 0.4, 0.5$
(Standard Errors Indicated in Parentheses)

λ	0.0	0.1	0.2	0.3	0.4	0.5
Regressor						
Intercept	5.277	4.986	4.640	4.220	3.703	3.05
	(0.393)	(0.469)	(0.566)	(0.694)	(0.868)	(1.12)
ln(FINCBTAX$_5$)	0.3719	0.4041	0.4425	0.4890	0.5463	0.6189
	(0.0432)	(0.0513)	(0.0619)	(0.0760)	(0.0954)	(0.123)
Age	0.0293	0.0283	0.0272	0.0259	0.02418	0.02205
	(0.0623)	(0.0621)	(0.0625)	(0.0640)	(0.0673)	(0.0738)
Age^2	-0.00034	-0.00033	-0.00031	-0.00029	-0.0271	-0.0243
	(0.00632)	(0.00638)	(0.00652)	(0.00680)	(0.00734)	(0.00829)
Any College	0.1994	0.1839	0.1655	0.1432	0.1157	0.0809
	(0.0359)	(0.0375)	(0.0401)	(0.0441)	(0.0506)	(0.0608)
Urban, Midwest	0.0653	0.0539	0.04043	0.0241	0.0040	-0.0215
	(0.0569)	(0.0568)	(0.0572)	(0.0581)	(0.0604)	(0.0651)
Urban, Not Midwest	0.1632	0.1497	0.1337	0.1144	0.0904	0.0602
	(0.0463)	(0.0478)	(0.0501)	(0.0538)	(0.0595)	(0.0688)
Renter	-0.1250	-0.1149	-0.1028	-0.0882	-0.0701	-0.0473
	(0.0146)	(0.0168)	(0.0198)	(0.0240)	(0.0299)	(0.0382)
Housing Assistance	-0.2378	-0.2086	-0.1738	-0.1317	-0.0080	-0.014
	(0.0690)	(0.0736)	(0.0808)	(0.0917)	(0.108)	(0.133)

missing-data issues only through weighting adjustment. Alternatives, e.g., single or multiple imputation, can also be considered. For some general discussion of missing-data adjustment methods, see, e.g., Little and Rubin (1987) and references cited therein. Rubin (1987) gives a detailed discussion of multiple imputation methods. Finally, it is worth noting that the statistical literature has tended to handle missing-data and errors-in-variables regression problems as separate issues. In light of the empirical results described here and in other literature, it is appropriate to focus additional attention on regression methods that account explicitly for both measurement errors and missing data.

Acknowledgements

The authors thank Fred Conrad for helpful comments on an earlier version of this paper. Some of this work was carried out when J. L. Eltinge and I. S. Yansaneh were visiting the Bureau of Labor Statistics under the ASA/NSF/BLS Research Fellow Program, and was supported by a grant from the National Science Foundation (SES-9022443). J. L. Eltinge's research was also supported in part by a grant from the National Institutes of Health (CA 57030). The views expressed here are those of the authors and do not necessarily represent the policies of the Bureau of Labor Statistics.

References

Browne, M.W. (1984). Asymptotically distribution-free methods in the analysis of covariance structures. Br. J. Math. Statist. Psychol., 37, 62-83.

Carroll, R. J. and Stefanski, L. A. (1990). Approximating Quasi-likelihood Estimation in Models with Surrogate Predictors. Journal of the American Statistical Association, 85, 652-663.

Czajka, J. L., Hirabayashi, S. M., Little, R. J. A. and Rubin, D. B. (1992). Projecting from Advance Data Using Propensity Modeling: An Application to Income and Tax Statistics. Journal of Business and Economic Statistics, 10, 117-131.

Eltinge, J. L. and Yansaneh, I. S. (1993). Weighting adjustments for income nonresponse in the U.S. Consumer Expenditure Survey. Technical Report #202, Department of Statistics, Texas A&M University.

Fuller, W. A. (1987). Measurement Error Models. New York: Wiley.

Fuller, W. A. (1991). Regression estimation in the presence of measurement error. In Measurement Errors in Surveys, ed. P. P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, and S. Sudman. New York: Wiley, 617-635.

Garner, T. I. and Blanciforti, L. A. (1994). Household Income Reporting: An Analysis of U. S. Consumer Expenditure Survey Data. Journal of Official Statistics, 10, in press.

Gillingham, R. and Hagemann, R. (1983). Cross-sectional estimation of a simultaneous model of tenure choice and housing services demand. Journal of Urban Economics, 14, 16-39.

Kalman, R.E. (1982). System identification from noisy data. In Dynamical Systems II, ed. A.R. Dednarek and L. Cesari. New York: Academic Press, 135-164.

Klepper, S. and Leamer, E.E. (1984). Consistent sets of estimates for regressions with errors in all variables. Econometrica, 52, 163-183.

Little, R. J. A. (1986). Survey nonresponse adjustments for estimates of means. International Statistical Review, 54, 139-157.

Little, R. J. A. (1991). Inference with survey weights. Journal of Official Statistics, 7, 405-424.

Little, R. J. A. and Rubin, D. B. (1987). Statistical Analysis with Missing Data. New York: Wiley.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. Biometrika, 70, 41-55.

Rubin, D. B. (1987). Multiple Imputation for Nonresponse in Surveys. New York: Wiley.

Sawtelle, B. A. (1993). Income elasticities of household expenditures: a U.S. cross-section perspective. Applied Economics, 25, 635-644.

Skinner, C. J. (1989). Domain means, regression and multivariate analysis. In Analysis of Complex Survey Data, ed. C. J. Skinner, D. Holt, and T. M. F. Smith. New York: Wiley, 59-87.

United States Bureau of Labor Statistics (1991). News: Consumer Expenditures in 1990. Publication USDL 91-607, United States Department of Labor, Washington, DC.

Wolter, K. M. (1985). Introduction to Variance Estimation. New York: Springer-Verlag.